UNCLASSIFIED

AD NUMBER
ADA194109
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to DoD only; Specific Authority; AUG 1987. Other requests shall be referred to Marine Corps Headquarters, Code RDS, Washington, DC.**
AUTHORITY
ONR 25 Apr 1983

THIS PAGE IS UNCLASSIFIED

(4)

CRM 87-151 / August 1987

DTIC FILE COPY

AD-A194 109

RESEARCH MEMORANDUM

DIMENSIONALITY OF THE MATH KNOWLEDGE ITEM POOL FOR THE ACCELERATED CAT-ASVAB PROJECT

D. R. Divgi

A Division of CNA Hudson Institute

CENTER FOR NAVAL ANALYSES

4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

88 4 25 090

Work conducted under contract N00014-87-C-0001.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
CRM 87-151

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
Center for Naval Analyses

6b. OFFICE SYMBOL (If applicable)
CNA

7a. NAME OF MONITORING ORGANIZATION
Commandant, of the Marine Corps (Code RDS)

6c. ADDRESS (City, State, and ZIP Code)
4401 Ford Avenue
Alexandria, Virginia 22302-0268

7b. ADDRESS (City, State, and ZIP Code)
Headquarters, Marine Corps
Washington, D.C. 20380

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
Office of Naval Research

8b. OFFICE SYMBOL (If applicable)
ONR

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
N00014-87-C-0001

8c. ADDRESS (City, State, and ZIP Code)
800 North Quincy Street
Arlington, Virginia 22217

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO
65153M	C0031		

11. TITLE (Include Security Classification)
Dimensionality of the Math Knowledge Item Pool for the Accelerated CAT-ASVAB Project

12. PERSONAL AUTHOR(S)
D. R. Divgi

13a. TYPE OF REPORT
Final

13b. TIME COVERED
FROM ________ TO ________

14. DATE OF REPORT (Year, Month, Day)
August 1987

15. PAGE COUNT
14

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
05	08	
12	03	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Aptitude tests, ASVAB (Armed Services Vocational Aptitude Battery), Bayes theorem, CAT (Computerized Adaptive Testing), Factor analysis, Statistical analysis, Test scores

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The theory underlying computerized adaptive tests assumes that all items for a given subtest measure a single dimension. This assumption was examined for the Math Knowledge items in the item pool developed for the Armed Services Vocational Aptitude Battery. Departures from the assumption were found to be minor.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
Major J. R. Robinson

22b. TELEPHONE (Include Area Code)
(703) 824-2643

22c. OFFICE SYMBOL
CNA

DD Form 1473, JUN 86 — Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

CNA

CENTER FOR NAVAL ANALYSES

A Division of Hudson Institute 4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268 • (703) 824-2000

22 September 1987

MEMORANDUM FOR DISTRIBUTION LIST

Subj: Center for Naval Analyses Research Memorandum 87-151

Encl: (1) CNA Research Memorandum 87-151, "Dimensionality of Math Knowledge Item Pool for the Accelerated CAT-ASVAB Project," by D. R. Divgi, August 1987

1. Enclosure (1) is forwarded as a matter of possible interest.

2. A computerized adaptive testing (CAT) version of the Armed Services Vocational Aptitude Battery (ASVAB) is being developed in the Accelerated CAT-ASVAB Project. The theory underlying CAT assumes that items in the pool for any subtest are unidimensional. This Research Memorandum examines the validity of the assumption for the Math Knowledge subtest by factor analyzing scores in different content areas.

William H. Sims
Director, Manpower and Training Program
Marine Corps Operations Analysis Group

Distribution List:
Reverse Page

Subj: Center for Naval Analyses Research Memorandum 87-151

Distribution List

SNDL

A1 ASSTSECNAV MRA
A1 DASN - MANPOWER (2 copies)
A6 HQMC MPR
  Attn: Code M
  Attn: Code MR
  Attn: Code MP
  Attn: Code MA
A6 HQMC TRNG
A6 HQMC RD&S
A6 HQMC RA
A6 HQMC AVN
E3D1 CNR
E3D5 NAVPERSRANDCEN
  Attn: Technical Director (Code 01)
  Attn: Technical Library
  Attn: CAT/ASVAB PMO
  Attn: Director Testing Systems (Code 63)
FF38 USNA
  Attn: Nimitz Library
FF42 NAVPGSCOL
FF44 NAVWARCOL
FJA1 COMNAVMILPERSCOM
FJB1 COMNAVCRUITCOM
FT1 CNET
V12 MCDEC
  Attn: Commanding General
  Attn: Development Center, Plans Division (Code D08) (2 copies)

OPNAV

OP-01
OP-11
OP-12
OP-13

OTHER

Joint Services CAT-ASVAB Working Group (16 copies)
Defense Advisory Committee on Military Personnel Testing (8 copies)

CRM 87-151 / August 1987

DIMENSIONALITY OF THE MATH KNOWLEDGE ITEM POOL FOR THE ACCELERATED CAT-ASVAB PROJECT

D. R. Divgi

Marine Corps Operations Analysis Group

A Division of CNA Hudson Institute

CENTER FOR NAVAL ANALYSES

4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

ABSTRACT

The theory underlying computerized adaptive tests assumes that all items for a given subtest measure a single dimension. This assumption was examined for the Math Knowledge items in the item pool developed for the Armed Services Vocational Aptitude Battery. Departures from the assumption were found to be minor.

EXECUTIVE SUMMARY

BACKGROUND

A computerized adaptive testing (CAT) version of the Armed Services Vocational Aptitude Battery (ASVAB) is being developed in the Accelerated CAT-ASVAB Project (ACAP). One fundamental assumption of item response theory (IRT), on which CAT-ASVAB is based, is that all items in the pool for a given subtest measure the same dimension, i.e., the same aptitude. If this assumption is violated, it may be necessary to impose content balancing, i.e., to ensure that the numbers of items in different content areas do not change from one examinee to another.

A factor analysis of item responses had shown the Math Knowledge subtest to be more troublesome than other subtests; while statistical tests indicated that there were four factors, the factors could not be given meaningful interpretations. The present study approaches the problem from a different perspective. Using a taxonomy provided by the Air Force Human Resources Laboratory, items in the Math Knowledge subtest were split into five content areas: fractions, decimals, and percents; analytic and plane/solid geometry; powers, exponents, and roots; equations and inequalities; and miscellaneous. Each item in the ACAP item pool was assigned to one of these areas by the author. The resulting scores in content areas, rather than individual items, were factor analyzed.

The data for the study consisted of item responses by applicants in the ACAP calibration sample. The item pool had been divided into five forms administered to equivalent random samples of applicants. The sample sizes for the forms ranged from 2,455 to 2,744.

METHODOLOGY

One analysis was performed using number-correct scores on the content areas. KR20 reliabilities of these scores were used as their communalities in the factor analyses.

The second analysis was carried out with ability estimates, using values of item parameters obtained during the IRT calibration of item pools. Again, estimated reliabilities were used as communalities.

RESULTS AND CONCLUSION

In the factor analysis of raw scores, a dominant first factor was obtained, although the second factor explained 4.7 percent to 11.4 percent of the variance in the five forms. However, the loadings on the second factor were related to the difficulty of the content area, indicating that the second factor was at least partly a spurious difficulty factor. Factor analysis of ability estimates showed the second factor to be much weaker, explaining 3.5 percent to 6.3 percent of the variance.

Thus, the evidence against unidimensionality is not strong enough to require content balancing in the Math Knowledge subtest of CAT-ASVAB.

TABLE OF CONTENTS

INTRODUCTION

Development of a computerized adaptive testing (CAT) version of the Armed Services Vocational Aptitude Battery (ASVAB) is being carried out in the Accelerated CAT-ASVAB Project (ACAP). CAT is based on a fundamental assumption of item response theory (IRT) that all items in the pool for a given subtest are unidimensional, i.e., they measure the same trait. The ACAP item pools were developed by Prestwood and Vale [1]. For estimating item parameters, the item pool for each subtest was divided into several forms; each form was administered to a sample of applicants. Dimensionalities of the item pools were examined by Segall and Moreno [2] using the TESTFACT program [3].

The Math Knowledge (MK) subtest turned out to be more troublesome than the others. While four factors were found to be statistically significant in each of five forms, Segall and Moreno were unable to interpret the factor solutions. One possible reason is that the data violated the assumption in TESTFACT that all abilities are normally distributed. Other assumptions may be invalid also.

The present study was carried out to analyze the data from a different perspective. The purpose of studying the dimensionality of an item pool is to decide if it is necessary to perform content balancing in CAT, as in the case of the General Science subtest [2]. In content balancing, each item is assigned to one category according to its content. Balancing consists of ensuring that the numbers of items administered from the various categories do not change from one examinee to another. The question is whether the traits measured by various categories differ enough to require content balancing. It can be answered by calculating separate scores on the content categories, and then factor analyzing these scores. No balancing is needed if the first factor is dominant and if the variance explained by the second factor is small.

Content analysis of ASVAB form 8a has resulted in a taxonomy of items in all subtests of the ASVAB (appendix A of [4]). Following a later version of this taxonomy, Math Knowledge items were divided into five content categories: fractions, decimals, and percents; analytic and plane/solid geometry; powers, exponents, and roots; equations and inequalities; and other, miscellaneous topics.

The data for the study consisted of item responses by applicants in Prestwood and Vale's calibration sample. The responses had already been scored as right, wrong, or unanswered. The sample sizes for the five forms ranged from 2,455 to 2,744.

METHODOLOGY

Prestwood and Vale's assignments of individual items to content categories are no longer available. Therefore the author used his own judgment to make these assignments. Number-correct scores on the five content areas were computed for each examinee and factor analyzed using the Statistical Analysis System [5]. Each form was analyzed separately.

The main purpose of the analysis was to identify the number of nontrivial factors. It is known that, when communalities are estimated from the correlations, one can always fit a single factor to three variables, no matter what the underlying reality is ([6], p. 138). This suggests that, especially when the number of variables is small, the number of factors is underestimated if communalities are fitted during the factor analysis. Therefore, the KR20 reliability of the score in each content area was calculated and used as its communality.

If the tests being factor analyzed differ appreciably in difficulty, one can obtain spurious "difficulty factors," especially if the scores are based on small numbers of items. The content areas in the Math Knowledge subtest often contained only four or five items. Therefore a second analysis was performed. Prestwood and Vale [1] have estimated IRT parameters for all items. These were used to calculate a Bayesian posterior mean ability estimate on each content area for each examinee. The ability estimates were factor analyzed, again using reliabilities as communalities. (Reliability was defined as the squared correlation with true ability. Regression of true ability on posterior mean is linear; variance of the means is the explained variance, and average posterior variance is the residual. Squared correlation equals the former divided by the sum of the two variances.)

In estimating abilities, unanswered items were treated as wrong. This approach was justified by the fact that the mean numbers of omitted and unreached items were only 0.22 and 0.04, respectively.

RESULTS

Table 1 shows that the numbers of items in the content areas vary substantially from one form to another. This was one of the reasons Segall and Moreno [2] could not interpret the results of their factor analyses of items. As expected, reliabilities go up and down with the number of items. On summing over forms, the total numbers of

TABLE 1

RESULTS FOR NUMBER-CORRECT SCORES

Content area	Number of items	Mean p-value	KR20	Factor loading: Factor 1	Factor 2
Form 1, N = 2,744, % of variance			=	86.5	10.7
Fractions	13	.719	.813	.772	.440
Geometry	4	.699	.486	.675	.136
Powers	5	.338	.804	.782	−.404
Equations	19	.524	.886	.926	−.045
Miscellaneous	5	.570	.628	.780	−.095
Form 2, N = 2,683, % of variance			=	92.5	4.7
Fractions	7	.662	.546	.686	.254
Geometry	5	.639	.361	.579	.109
Powers	11	.521	.837	.876	−.250
Equations	15	.586	.843	.901	−.059
Miscellaneous	8	.480	.494	.683	.050
Form 3, N = 2,649, % of variance			=	90.7	5.3
Fractions	8	.721	.662	.740	.278
Geometry	8	.521	.594	.717	.103
Powers	10	.464	.840	.853	−.304
Equations	15	.545	.875	.918	−.068
Miscellaneous	5	.713	.549	.749	.056
Form 4, N = 2,540, % of variance			=	92.9	6.6
Fractions	13	.717	.764	.815	.295
Geometry	6	.550	.402	.639	.009
Powers	8	.462	.805	.830	−.322
Equations	14	.555	.819	.899	−.062
Miscellaneous	5	.718	.382	.614	.125
Form 5, N = 2,455, % of variance			=	82.1	11.4
Fractions	7	.651	.614	.729	.037
Geometry	7	.602	.552	.701	−.026
Powers	9	.464	.836	.849	−.264
Equations	18	.525	.859	.908	−.076
Miscellaneous	5	.921	.582	.508	.560

items in the five areas differ from those reported by Prestwood and Vale ([1], table 25). The differences result from the unavoidable subjectivity in judging item content.

Two principal factors were extracted for each form. The percentage of variance explained by the second factor, as reported in table 1, varied from 4.7 to 11.4. These values imply that the second factor is not small enough to be ignored. However, loadings on the second factor are related to easiness of the content area as expressed in the mean p-value, being generally larger for easier than for harder content areas. This relationship may indicate that there are two distinct factors, "lower math" and "higher math," that can be identified by oblique rotation of the principal factors. However, it is also possible that the second factor is partly a spurious difficulty factor, resulting from nonlinear relationships among content-area scores.

Results in table 2 support the second interpretation. When ability estimates rather than raw scores are analyzed, the second factor explains 3.5 percent to 6.2 percent of the variance and thus appears weak enough to be ignored. Although the loadings show the same relationship to easiness of the content area as before, they are so small that an oblique rotation will not yield clearly distinct factors. "Lower math" contains fractions and geometry; "higher math" includes powers and equations, both of which involve algebraic symbols. However, the separation between these two factors is comparable with that between variables belonging to the same factor. This result is not strong enough to require content balancing in CAT-ASVAB.

Factor scores computed from ability estimates were analyzed to see if their distributions were non-normal. The results were consistent from one form to another. Distributions for the second factor were symmetric. Those for the first factor were positively skewed, the skewness ranging from 0.21 to 0.36. They were also short-tailed, the kurtosis ranging from -0.68 to -0.98. These statistics suggest that the assumption of normality, required by TESTFACT [3], was violated.

CONCLUSION

In view of the results in table 2, the evidence for two or more dimensions in the Math Knowledge item pool is weak, and therefore content balancing is not necessary.

TABLE 2

FACTOR ANALYSIS OF ABILITY ESTIMATES

Content area	Reliability		Factor loading: Factor 1	Factor loading: Factor 2
Form 1, % of variance		=	92.5	6.2
Fractions	.768		.802	.341
Geometry	.526		.700	.106
Powers	.642		.772	−.259
Equations	.871		.930	−.038
Miscellaneous	.621		.793	−.143
Form 2, % of variance		=	93.6	4.1
Fractions	.592		.716	.263
Geometry	.376		.600	.070
Powers	.816		.874	−.204
Equations	.861		.912	−.113
Miscellaneous	.585		.743	.069
Form 3, % of variance		=	94.6	3.5
Fractions	.645		.763	.057
Geometry	.641		.749	.257
Powers	.785		.860	−.190
Equations	.862		.923	−.120
Miscellaneous	.576		.764	.049
Form 4, % of variance		=	94.7	3.6
Fractions	.759		.842	.179
Geometry	.523		.702	−.044
Powers	.747		.834	−.217
Equations	.827		.911	−.071
Miscellaneous	.516		.683	.185
Form 5, % of variance		=	91.7	5.0
Fractions	.610		.753	.047
Geometry	.592		.742	−.059
Powers	.806		.872	−.184
Equations	.863		.920	−.012
Miscellaneous	.394		.517	.350

REFERENCES

[1] Air Force Human Resources Laboratory TR-85-19, *Armed Services Vocational Aptitude Battery: Development of an Adaptive Item Pool*, by J. Stephen Prestwood, C. David Vale, Randy H. Massey, and John R. Welsh, Sep 1985

[2] Naval Postgraduate School Memorandum, "Minutes of March 1986 Meeting of the CAT-ASVAB Psychometric Committee," by Bruce Bloxom, 23 May 1986

[3] National Opinion Research Center Report 85-1, *Full-Information Item Factor Analysis*, by R. Darrell Bock, Robert Gibbons, and Eiji Muraki, revised, July 1986

[4] U.S. Military Entrance Processing Command. *Technical Supplement to the Counselor's Manual for the Armed Services Vocational Aptitude Battery Form 14*. North Chicago: U.S. Military Entrance Processing Command, 1985

[5] SAS Institute Inc. *Statistical Analysis System, Version 5.03*. Cary, NC: SAS Institute, Inc., 1986

[6] Stanley A. Mulaik, *The Foundations of Factor Analysis*. New York: McGraw-Hill, 1972